

/600



ENTERED

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/586,106D

DATE: 02/12/2003 P.6
TIME: 12:34:45

Input Set : A:\08411-032001.txt

Output Set: N:\CRF4\02122003\I586106D.raw

```

4 <110> APPLICANT: Wright, David A.
5   Voytas, Daniel F.
7 <120> TITLE OF INVENTION: PLANT RETROELEMENTS AND METHODS RELATED THERETO
9 <130> FILE REFERENCE: P-1065A
11 <140> CURRENT APPLICATION NUMBER: 09/586,106D
C--> 12 <141> CURRENT FILING DATE: 2003-02-07
14 <150> PRIOR APPLICATION NUMBER: 60/087,125
15 <151> PRIOR FILING DATE: 1998-05-29
17 <150> PRIOR APPLICATION NUMBER: 09/322,478
18 <151> PRIOR FILING DATE: 1999-05-28
20 <160> NUMBER OF SEQ ID NOS: 190
22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 18
26 <212> TYPE: DNA
27 <213> ORGANISM: Glycine max
29 <400> SEQUENCE: 1
30 tggcgccggtt gccaatg                                     18
32 <210> SEQ ID NO: 2
33 <211> LENGTH: 18
34 <212> TYPE: DNA
35 <213> ORGANISM: Glycine max
37 <400> SEQUENCE: 2
38 tggcgccggtt gtcgggga                                     18
40 <210> SEQ ID NO: 3
41 <211> LENGTH: 6
42 <212> TYPE: DNA
43 <213> ORGANISM: Glycine max
45 <400> SEQUENCE: 3
46 ttgggg                                                  6
48 <210> SEQ ID NO: 4
49 <211> LENGTH: 7
50 <212> TYPE: PRT
51 <213> ORGANISM: Artificial Sequence
53 <220> FEATURE:
54 <223> OTHER INFORMATION: plant retroelement sequence
56 <400> SEQUENCE: 4
57 Met Ala Ser Arg Lys Arg Lys
58 1               5
60 <210> SEQ ID NO: 5
61 <211> LENGTH: 1263
62 <212> TYPE: DNA
63 <213> ORGANISM: Artificial Sequence

```

RAW SEQUENCE LISTING

DATE: 02/12/2003

PATENT APPLICATION: US/09/586,106D

TIME: 12:34:45

Input Set : A:\08411-032001.txt

Output Set: N:\CRF4\02122003\I586106D.raw

65 <220> FEATURE:

66 <223> OTHER INFORMATION: plant retroelement sequence

68 <400> SEQUENCE: 5

```

69 atggcctccc gtaaacgcaa agctgtgccc acacccgggg aagcgtccaa ctgggactct      60
70 tcacgtttca ctttcgagat tgcttggcac agataccagg atagcattca gctccggaac      120
71 atccttccag agaggaatgt agagcttgga ccagggatgt ttgatgagtt cctgcaggaa      180
72 ctccagaggc tcagatggga ccaggttctg acccgacttc cagagaagtg gattgatgtt      240
73 gctctggtga aggagtttta ctccaaccta tatgatccag aggaccacag tccgaagttt      300
74 tggagtgttc gaggacaggt tgtgagattt gatgctgaga cgattaatga tttcctcgac      360
75 accccggtca tcttggcaga gggagaggat tatccagcct actctcagta cctcagcact      420
76 cctccagacc atgatgccat cctttccgct ctgtgtactc cagggggacg atttgttctg      480
77 aatgttgata gtgccccctg gaagctgctg cggaaggatc tgatgacgct cgcgagaca      540
78 tggagtgtgc tctcttattt taaccttgca ctgacttttc acacttctga tattaatgtt      600
79 gacagggccc gactcaatta tggcttggtg atgaagatgg acctggacgt gggcagcctc      660
80 atttctcttc agatcagtc gatcgcccag tccatcactt ccaggcttgg gttcccagcg      720
81 ttgatcacia cactgtgtga gattcagggg gttgtctctg ataccctgat ttttgagtca      780
82 ctcatgctct tgatcaacct tgcctacatt aagaagaact gctggaaccc tgccgatcca      840
83 tctatcacat ttcaggggac ccgccgcacg cgcaccagag cttcggcgct ggcatctgag      900
84 gctcctcttc catcccagca tccttctcag cctttttccc agagaccacg gcctccactt      960
85 ctatccacct cagcacctcc atacatgcat ggacagatgc tcaggtcctt gtaccagggt      1020
86 cagcagatca tcattcagaa cctgtatcga ttgtccctac atttgagat ggatctgcca      1080
87 ctcatgactc cggaggccta tcgtcagcag gtcgccaagc taggagacca gccctccact      1140
88 gacagggggg aagagccttc tggagccgct gctactgagg atcctgccgt tgatgaagac      1200
89 ctcatagctg acttggctgg cgctgattgg agcccatggg cagacttggg cagaggcagc      1260
90 tga                                                                 1263

```

92 <210> SEQ ID NO: 6

93 <211> LENGTH: 421

94 <212> TYPE: PRT

95 <213> ORGANISM: Artificial Sequence

97 <220> FEATURE:

98 <223> OTHER INFORMATION: plant retroelement sequence

100 <400> SEQUENCE: 6

```

101 Met Ala Ser Arg Lys Arg Lys Ala Val Pro Thr Pro Gly Glu Ala Ser
102 1 5 10 15
103 Asn Trp Asp Ser Ser Arg Phe Thr Phe Glu Ile Ala Trp His Arg Tyr
104 20 25 30
105 Gln Asp Ser Ile Gln Leu Arg Asn Ile Leu Pro Glu Arg Asn Val Glu
106 35 40 45
107 Leu Gly Pro Gly Met Phe Asp Glu Phe Leu Gln Glu Leu Gln Arg Leu
108 50 55 60
109 Arg Trp Asp Gln Val Leu Thr Arg Leu Pro Glu Lys Trp Ile Asp Val
110 65 70 75 80
111 Ala Leu Val Lys Glu Phe Tyr Ser Asn Leu Tyr Asp Pro Glu Asp His
112 85 90 95
113 Ser Pro Lys Phe Trp Ser Val Arg Gly Gln Val Val Arg Phe Asp Ala
114 100 105 110
115 Glu Thr Ile Asn Asp Phe Leu Asp Thr Pro Val Ile Leu Ala Glu Gly
116 115 120 125
117 Glu Asp Tyr Pro Ala Tyr Ser Gln Tyr Leu Ser Thr Pro Pro Asp His

```

RAW SEQUENCE LISTING

DATE: 02/12/2003

PATENT APPLICATION: US/09/586,106D

TIME: 12:34:45

Input Set : A:\08411-032001.txt

Output Set: N:\CRF4\02122003\I586106D.raw

```

118      130      135      140
119 Asp Ala Ile Leu Ser Ala Leu Cys Thr Pro Gly Gly Arg Phe Val Leu
120 145      150      155      160
121 Asn Val Asp Ser Ala Pro Trp Lys Leu Leu Arg Lys Asp Leu Met Thr
122      165      170      175
123 Leu Ala Gln Thr Trp Ser Val Leu Ser Tyr Phe Asn Leu Ala Leu Thr
124      180      185      190
125 Phe His Thr Ser Asp Ile Asn Val Asp Arg Ala Arg Leu Asn Tyr Gly
126      195      200      205
127 Leu Val Met Lys Met Asp Leu Asp Val Gly Ser Leu Ile Ser Leu Gln
128      210      215      220
129 Ile Ser Gln Ile Ala Gln Ser Ile Thr Ser Arg Leu Gly Phe Pro Ala
130 225      230      235      240
131 Leu Ile Thr Thr Leu Cys Glu Ile Gln Gly Val Val Ser Asp Thr Leu
132      245      250      255
133 Ile Phe Glu Ser Leu Ser Pro Val Ile Asn Leu Ala Tyr Ile Lys Lys
134      260      265      270
135 Asn Cys Trp Asn Pro Ala Asp Pro Ser Ile Thr Phe Gln Gly Thr Arg
136      275      280      285
137 Arg Thr Arg Thr Arg Ala Ser Ala Ser Ala Ser Glu Ala Pro Leu Pro
138      290      295      300
139 Ser Gln His Pro Ser Gln Pro Phe Ser Gln Arg Pro Arg Pro Pro Leu
140 305      310      315      320
141 Leu Ser Thr Ser Ala Pro Pro Tyr Met His Gly Gln Met Leu Arg Ser
142      325      330      335
143 Leu Tyr Gln Gly Gln Gln Ile Ile Ile Gln Asn Leu Tyr Arg Leu Ser
144      340      345      350
145 Leu His Leu Gln Met Asp Leu Pro Leu Met Thr Pro Glu Ala Tyr Arg
146      355      360      365
147 Gln Gln Val Ala Lys Leu Gly Asp Gln Pro Ser Thr Asp Arg Gly Glu
148      370      375      380
149 Glu Pro Ser Gly Ala Ala Ala Thr Glu Asp Pro Ala Val Asp Glu Asp
150 385      390      395      400
151 Leu Ile Ala Asp Leu Ala Gly Ala Asp Trp Ser Pro Trp Ala Asp Leu
152      405      410      415
153 Gly Arg Gly Ser Glx
154      420

```

156 <210> SEQ ID NO: 7

157 <211> LENGTH: 1596

158 <212> TYPE: DNA

159 <213> ORGANISM: Artificial Sequence

161 <220> FEATURE:

162 <223> OTHER INFORMATION: plant retroelement sequence

164 <400> SEQUENCE: 7

```

165 atgcgaggtga gaactgcattc tggagacgtt gttcctatta acttagaaat tgaagctacg      60
166 tgtcggcgta acaacgctgc aagaagaaga agggagcaag acatagaagg aagtagttac      120
167 acctcacctc ctccttctcc aaattatgct cagatggacg gggaaccggc acaaagagtc      180
168 aactagagg acttctctaa taccaccact cctcagttct ttacaagtat cacaaggccg      240
169 gaagtccaag cagatctcct tactcaaggg aacctcttcc atggtcttcc aaatgaagat      300

```

RAW SEQUENCE LISTING

DATE: 02/12/2003

PATENT APPLICATION: US/09/586,106D

TIME: 12:34:45

Input Set : A:\08411-032001.txt

Output Set: N:\CRF4\02122003\I586106D.raw

```

170 ccatatgcgc atctagcctc atacatagag atatgcagca ccgttaaaat cgccggagtt 360
171 ccaaaagatg cgatactcct taacctcttt tccttttccc tagcaggaga ggcaaaaaga 420
172 tggttgcact ccttttaaagg caatagctta agaacatggg aagaagtagt ggaaaaattc 480
173 ttaaagaagt atttcccaga gtcaaagacc gtcgaacgaa agatggagat ttcttatttc 540
174 catcaatttc tggatgaatc ccttagcgaa gcactagacc atttccacgg attgctaaga 600
175 aaaacaccaa cacacagata cagcgagcca gtacaactaa acatattcat cgatgacttg 660
176 caactcttaa tcgaaacagc tactagaggg aagatcaagc tgaagactcc cgaagaagcg 720
177 atggagctcg tcgagaacat ggcggctagc gatcaagcaa tccttcatga tcacacttat 780
178 gttcccacaa aaagaagcct cttggagctt agcacgcagg acgcaacttt ggtacaaaac 840
179 aagctggtga cgaggcagat agaagccctc atcgaaaccc tcagcaagct gcctcaacaa 900
180 ttacaagcga taagtctctc ccactcttct gttttgcagg tagaagaatg ccccatatgc 960
181 agagggacac atgagcctgg acaatgtgca agccaacaag acccctctcg tgaagtaaata 1020
182 tatataggca tactaaatcg ttacggattt cagggctaca accagggaaa tccatctgga 1080
183 ttcaatcaag gggcaacaag atttaatcac gagccaccgg ggtttaatca aggaagaaac 1140
184 ttcatgcaag gctcaagttg gacgaataaa ggaaatcaat ataaggagca aaggaaccaa 1200
185 ccaccatacc agccaccata ccagcaccct agccaaggtc cgaatcagca agaaaagccc 1260
186 accaaaatag aggaactgct gctgcaattc atcaaggaga caagatcaca tcaaaagagc 1320
187 acggatgcag ccattcggaa tctagaagtt caaatgggccc aactggcgca tgacaaagcc 1380
188 gaacggccca ctagaacttt cgggtgctaac atggagagaa gaaccccaag gaaggataaa 1440
189 gcagtactga ctagagggca gagaagagcg caggaggagg gtaagggtga aggagaagac 1500
190 tggccagaag aaggaaggac agagaagaca gaagaagaag agaaggtggc agaagaacct 1560
191 aagcgtacca agagccagag agcaagggaa gccaaag 1596

```

193 <210> SEQ ID NO: 8

194 <211> LENGTH: 532

195 <212> TYPE: PRT

196 <213> ORGANISM: Artificial Sequence

198 <220> FEATURE:

199 <223> OTHER INFORMATION: plant retroelement sequence

201 <400> SEQUENCE: 8

```

202 Met Arg Gly Arg Thr Ala Ser Gly Asp Val Val Pro Ile Asn Leu Glu
203 1 5 10 15
204 Ile Glu Ala Thr Cys Arg Arg Asn Asn Ala Ala Arg Arg Arg Arg Glu
205 20 25 30
206 Gln Asp Ile Glu Gly Ser Ser Tyr Thr Ser Pro Pro Pro Ser Pro Asn
207 35 40 45
208 Tyr Ala Gln Met Asp Gly Glu Pro Ala Gln Arg Val Thr Leu Glu Asp
209 50 55 60
210 Phe Ser Asn Thr Thr Thr Pro Gln Phe Phe Thr Ser Ile Thr Arg Pro
211 65 70 75 80
212 Glu Val Gln Ala Asp Leu Leu Thr Gln Gly Asn Leu Phe His Gly Leu
213 85 90 95
214 Pro Asn Glu Asp Pro Tyr Ala His Leu Ala Ser Tyr Ile Glu Ile Cys
215 100 105 110
216 Ser Thr Val Lys Ile Ala Gly Val Pro Lys Asp Ala Ile Leu Leu Asn
217 115 120 125
218 Leu Phe Ser Phe Ser Leu Ala Gly Glu Ala Lys Arg Trp Leu His Ser
219 130 135 140
220 Phe Lys Gly Asn Ser Leu Arg Thr Trp Glu Glu Val Val Glu Lys Phe
221 145 150 155 160

```

RAW SEQUENCE LISTING

DATE: 02/12/2003

PATENT APPLICATION: US/09/586,106D

TIME: 12:34:45

Input Set : A:\08411-032001.txt

Output Set: N:\CRF4\02122003\I586106D.raw

```

222 Leu Lys Lys Tyr Phe Pro Glu Ser Lys Thr Val Glu Arg Lys Met Glu
223      165      170      175
224 Ile Ser Tyr Phe His Gln Phe Leu Asp Glu Ser Leu Ser Glu Ala Leu
225      180      185      190
226 Asp His Phe His Gly Leu Leu Arg Lys Thr Pro Thr His Arg Tyr Ser
227      195      200      205
228 Glu Pro Val Gln Leu Asn Ile Phe Ile Asp Asp Leu Gln Leu Leu Ile
229      210      215      220
230 Glu Thr Ala Thr Arg Gly Lys Ile Lys Leu Lys Thr Pro Glu Glu Ala
231 225      230      235      240
232 Met Glu Leu Val Glu Asn Met Ala Ala Ser Asp Gln Ala Ile Leu His
233      245      250      255
234 Asp His Thr Tyr Val Pro Thr Lys Arg Ser Leu Leu Glu Leu Ser Thr
235      260      265      270
236 Gln Asp Ala Thr Leu Val Gln Asn Lys Leu Leu Thr Arg Gln Ile Glu
237      275      280      285
238 Ala Leu Ile Glu Thr Leu Ser Lys Leu Pro Gln Gln Leu Gln Ala Ile
239      290      295      300
240 Ser Ser Ser His Ser Ser Val Leu Gln Val Glu Glu Cys Pro Thr Cys
241 305      310      315      320
242 Arg Gly Thr His Glu Pro Gly Gln Cys Ala Ser Gln Gln Asp Pro Ser
243      325      330      335
244 Arg Glu Val Asn Tyr Ile Gly Ile Leu Asn Arg Tyr Gly Phe Gln Gly
245      340      345      350
246 Tyr Asn Gln Gly Asn Pro Ser Gly Phe Asn Gln Gly Ala Thr Arg Phe
247      355      360      365
248 Asn His Glu Pro Pro Gly Phe Asn Gln Gly Arg Asn Phe Met Gln Gly
249      370      375      380
250 Ser Ser Trp Thr Asn Lys Gly Asn Gln Tyr Lys Glu Gln Arg Asn Gln
251 385      390      395      400
252 Pro Pro Tyr Gln Pro Pro Tyr Gln His Pro Ser Gln Gly Pro Asn Gln
253      405      410      415
254 Gln Glu Lys Pro Thr Lys Ile Glu Glu Leu Leu Leu Gln Phe Ile Lys
255      420      425      430
256 Glu Thr Arg Ser His Gln Lys Ser Thr Asp Ala Ala Ile Arg Asn Leu
257      435      440      445
258 Glu Val Gln Met Gly Gln Leu Ala His Asp Lys Ala Glu Arg Pro Thr
259      450      455      460
260 Arg Thr Phe Gly Ala Asn Met Glu Arg Arg Thr Pro Arg Lys Asp Lys
261 465      470      475      480
262 Ala Val Leu Thr Arg Gly Gln Arg Arg Ala Gln Glu Glu Gly Lys Val
263      485      490      495
264 Glu Gly Glu Asp Trp Pro Glu Glu Gly Arg Thr Glu Lys Thr Glu Glu
265      500      505      510
266 Glu Glu Lys Val Ala Glu Glu Pro Lys Arg Thr Lys Ser Gln Arg Ala
267      515      520      525
268 Arg Glu Ala Lys
269      530
271 <210> SEQ ID NO: 9

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/586,106D

DATE: 02/12/2003
TIME: 12:34:46

Input Set : A:\08411-032001.txt
Output Set: N:\CRF4\02122003\I586106D.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:166; N Pos. 6,15,16,18
Seq#:168; N Pos. 7
Seq#:178; Xaa Pos. 2,3,5,6,7,9,10,11,12
Seq#:181; Xaa Pos. 2
Seq#:182; Xaa Pos. 1,9,10
Seq#:183; Xaa Pos. 12
Seq#:184; Xaa Pos. 4,8,13,15,19,23,26
Seq#:185; Xaa Pos. 4,6,7,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27
Seq#:185; Xaa Pos. 28,29,30,31,32,33,34,35,36,37,38,39,40,43
Seq#:186; Xaa Pos. 6
Seq#:187; Xaa Pos. 7,10
Seq#:188; N Pos. 11,12
Seq#:189; Xaa Pos. 5,8

VERIFICATION SUMMARY

DATE: 02/12/2003

PATENT APPLICATION: US/09/586,106D

TIME: 12:34:46

Input Set : A:\08411-032001.txt

Output Set: N:\CRF4\02122003\I586106D.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:5702 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5706 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:166
L:5707 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:166 after pos.:0
L:5729 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5733 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:168
L:5734 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:168 after pos.:0
L:5841 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5845 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:178
L:5846 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:178 after pos.:0
L:5881 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5885 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:181
L:5886 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:181 after pos.:0
L:5897 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5901 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:182
L:5905 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:182
L:5909 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:182
L:5910 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:182 after pos.:0
L:5921 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5925 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:183
L:5926 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:183 after pos.:0
L:5939 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5943 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:184
L:5947 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:184
L:5951 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:184
L:5955 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:184
L:5959 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:184
L:5963 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:184
L:5967 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:184
L:5968 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:184 after pos.:0
M:341 Repeated in SeqNo=184
L:5981 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:5985 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:185
L:5986 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:185 after pos.:0
M:341 Repeated in SeqNo=185
L:6001 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:6005 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:186
L:6006 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:186 after pos.:0
L:6017 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:6021 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:187
L:6025 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:187
L:6026 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:187 after pos.:0
L:6037 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:6041 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:188
L:6042 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:188 after pos.:0
L:6052 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:6056 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:189
L:6057 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:189 after pos.:0